Project 1

Problem 1

We will consider the "Boston" housing data set, from the "MASS" library in R. More details can be found on page 201 of the textbook of James, Witten, Hastie, and Tibshirani

1. Based on this data set, provide an estimate for the population coefficient of variation, σ/μ, of "medv". Call this estimate $\hat{\sigma}/\hat{\mu}$.

2. Propose an algorithm to obtain the standard normal bootstrap confidence interval. Use the algorithm proposed to compute a 95% standard normal bootstrap confidence interval for the population coefficient of variation, σ/μ, of "medv".

3. Propose an algorithm to obtain a bootstrap t confidence interval. Use the algorithm proposed to compute a bootstrap t confidence interval for the population coefficient of variation, σ/μ, of "medv" with confidence level 95%.

4. Propose an algorithm to obtain a BC_a bootstrap confidence interval. Use the algorithm proposed to compute a 95% BC_a bootstrap confidence interval for the population coefficient of variation, σ/μ, of "medv".

Problem 2

The goal of problem 2 is to understand cross validation.

1. Read pages 175 –183 of the textbook of James, Witten, Hastie, and Tibshirani and give a short summary.

2. Propose an algorithm for the $k-$ fold cross validation.

3. Propose an algorithm for the LOOCV.

4. In example 7.18 on page 210 of the textbook of Rizzo, LOOCV was used to select the best fitting model. Repeat the same analysis using a 10-fold cross validation. Compare the results using the two methods.

Problem 3

Solve problem 7.8 on page 213 of the textbook of Rizzo.